

## RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.

Application Serial Number: 10/541,513  
Source: PT/10  
Date Processed by STIC: 7/19/05

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 07/19/2005

PATENT APPLICATION: US/10/541,513

TIME: 15:11:28

Input Set : A:\Final sequence list-12810-00106-US.txt

Output Set: N:\CRF4\07192005\J541513.raw

3 <110> APPLICANT: Steiger, Sabine  
 4 Sandmann, Gerhard  
 6 <120> TITLE OF INVENTION: Method for producing ketocarotenoids by cultivating  
 genetically  
 7 modified organisms  
 9 <130> FILE REFERENCE: 12810-00106-US  
 C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/541,513  
 C--> 11 <141> CURRENT FILING DATE: 2005-07-08  
 11 <150> PRIOR APPLICATION NUMBER: PCT/EP2003/014876  
 12 <151> PRIOR FILING DATE: 2003-12-24  
 14 <150> PRIOR APPLICATION NUMBER: DE 103 00 649.4  
 15 <151> PRIOR FILING DATE: 2003-01-09  
 17 <160> NUMBER OF SEQ ID NOS: 15  
 19 <170> SOFTWARE: PatentIn version 3.3  
 22 <210> SEQ ID NO: 1  
 23 <211> LENGTH: 789  
 24 <212> TYPE: DNA  
 25 <213> ORGANISM: Nostoc sp. PCC73102  
 27 <220> FEATURE:  
 28 <221> NAME/KEY: CDS  
 29 <222> LOCATION: (1)..(789)  
 31 <400> SEQUENCE: 1  
 32 ttg aat ttt tgt gat aaa cca gtt agc tat tat gtt gca ata gag caa 48  
 33 Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln  
 34 1 5 10 15  
 36 tta agt gct aaa gaa gat act gtt tgg ggg ctg gtg att gtc ata gta 96  
 37 Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val  
 38 20 25 30  
 40 att att agt ctt tgg gta gct agt ttg gct ttt tta cta gct att aat 144  
 41 Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn  
 42 35 40 45  
 44 tat gcc aaa gtc cca att tgg ttg ata cct att gca ata gtt tgg caa 192  
 45 Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln  
 46 50 55 60  
 48 atg ttc ctt tat aca ggg cta ttt att act gca cat gat gct atg cat 240  
 49 Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His  
 50 65 70 75 80  
 52 ggg tca gtt tat cgt aaa aat ccc aaa att aat aat ttt atc ggt tca 288  
 53 Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser  
 54 85 90 95  
 56 cta gct gta gcg ctt tac gct gtg ttt cca tat caa cag atg tta aag 336  
 57 Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys  
 58 100 105 110  
 60 aat cat tgc tta cat cat cgt cat cct gct agc gaa gtt gac cca gat 384

## RAW SEQUENCE LISTING

DATE: 07/19/2005

PATENT APPLICATION: US/10/541,513

TIME: 15:11:28

Input Set : A:\Final sequence list-12810-00106-US.txt

Output Set: N:\CRF4\07192005\J541513.raw

```

61 Asn His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp
62      115      120      125
64 ttt cat gat ggt aag aga aca aac gct att ttc tgg tat ctc cat ttc      432
65 Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe
66      130      135      140
68 atg ata gaa tac tcc agt tgg caa cag tta ata gta cta act atc cta      480
69 Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu
70 145      150      155      160
72 ttt aat tta gct aaa tac gtt ttg cac atc cat caa ata aat ctc atc      528
73 Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile
74      165      170      175
76 tta ttt tgg agt att cct cca att tta agt tcc att caa ctg ttt tat      576
77 Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr
78      180      185      190
80 ttc gga aca ttt ttg cct cat cga gaa ccc aag aaa gga tat gtt tat      624
81 Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr
82      195      200      205
84 ccc cat tgc agc caa aca ata aaa ttg cca act ttt ttg tca ttt atc      672
85 Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile
86      210      215      220
88 gct tgc tac cac ttt ggt tat cat gaa gaa cat cat gag tat ccc cat      720
89 Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
90 225      230      235      240
92 gta cct tgg tgg caa ctt cca tct gta tat aag cag aga gta ttc aac      768
93 Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn
94      245      250      255
96 aat tca gta acc aat tcg taa      789
97 Asn Ser Val Thr Asn Ser
98      260
101 <210> SEQ ID NO: 2
102 <211> LENGTH: 262
103 <212> TYPE: PRT
104 <213> ORGANISM: Nostoc sp. PCC73102
107 <400> SEQUENCE: 2
109 Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln
110 1      5      10      15
113 Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val
114      20      25      30
117 Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn
118      35      40      45
121 Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln
122      50      55      60
125 Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His
126 65      70      75      80
129 Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser
130      85      90      95
133 Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys
134      100      105      110
137 Asn His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp

```

## RAW SEQUENCE LISTING

DATE: 07/19/2005

PATENT APPLICATION: US/10/541,513

TIME: 15:11:28

Input Set : A:\Final sequence list-12810-00106-US.txt

Output Set: N:\CRF4\07192005\J541513.raw

```

138          115          120          125
141 Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe
142          130          135          140
145 Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu
146 145          150          155          160
149 Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile
150          165          170          175
153 Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr
154          180          185          190
157 Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr
158          195          200          205
161 Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile
162          210          215          220
165 Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
166 225          230          235          240
169 Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn
170          245          250          255
173 Asn Ser Val Thr Asn Ser
174          260
177 <210> SEQ ID NO: 3
178 <211> LENGTH: 762
179 <212> TYPE: DNA
180 <213> ORGANISM: Nostoc sp. PCC73102
182 <220> FEATURE:
183 <221> NAME/KEY: CDS
184 <222> LOCATION: (1)..(762)
186 <400> SEQUENCE: 3
187 gtg atc cag tta gaa caa cca ctc agt cat caa gca aaa ctg act cca      48
188 Val Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro
189 1          5          10          15
191 gta ctg aga agt aaa tct cag ttt aag ggg ctt ttc att gct att gtc      96
192 Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val
193          20          25          30
195 att gtt agc gca tgg gtc att agc ctg agt tta tta ctt tcc ctt gac      144
196 Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Leu Ser Leu Asp
197          35          40          45
199 atc tca aag cta aaa ttt tgg atg tta ttg cct gtt ata cta tgg caa      192
200 Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln
201          50          55          60
203 aca ttt tta tat acg gga tta ttt att aca tct cat gat gcc atg cat      240
204 Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His
205 65          70          75          80
207 ggc gta gta ttt ccc caa aac acc aag att aat cat ttg att gga aca      288
208 Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr
209          85          90          95
211 ttg acc cta tcc ctt tat ggt ctt tta cca tat caa aaa cta ttg aaa      336
212 Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys
213          100          105          110
215 aaa cat tgg tta cac cac cac aat cca gca agc tca ata gac ccg gat      384

```

## RAW SEQUENCE LISTING

DATE: 07/19/2005

PATENT APPLICATION: US/10/541,513

TIME: 15:11:28

Input Set : A:\Final sequence list-12810-00106-US.txt

Output Set: N:\CRF4\07192005\J541513.raw

```

216 Lys His Trp Leu His His His Asn Pro Ala Ser Ser Ile Asp Pro Asp
217      115      120      125
219 ttt cac aat ggt aaa cac caa agt ttc ttt gct tgg tat ttt cat ttt      432
220 Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe
221      130      135      140
223 atg aaa ggt tac tgg agt tgg ggg caa ata att gcg ttg act att att      480
224 Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile
225 145      150      155      160
227 tat aac ttt gct aaa tac ata ctc cat atc cca agt gat aat cta act      528
228 Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr
229      165      170      175
231 tac ttt tgg gtg cta ccc tcg ctt tta agt tca tta caa tta ttc tat      576
232 Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr
233      180      185      190
235 ttt ggt act ttt tta ccc cat agt gaa cca ata ggg ggt tat gtt cag      624
236 Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln
237      195      200      205
239 cct cat tgt gcc caa aca att agc cgt cct att tgg tgg tca ttt atc      672
240 Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile
241      210      215      220
243 acg tgc tat cat ttt ggc tac cac gag gaa cat cac gaa tat cct cat      720
244 Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
245 225      230      235      240
247 att tct tgg tgg cag tta cca gaa att tac aaa gca aaa tag      762
248 Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys
249      245      250
252 <210> SEQ ID NO: 4
253 <211> LENGTH: 253
254 <212> TYPE: PRT
255 <213> ORGANISM: Nostoc sp. PCC73102
257 <400> SEQUENCE: 4
259 Val Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro
260 1      5      10      15
263 Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val
264      20      25      30
267 Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Leu Ser Leu Asp
268      35      40      45
271 Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln
272      50      55      60
275 Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His
276 65      70      75      80
279 Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr
280      85      90      95
283 Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys
284      100      105      110
287 Lys His Trp Leu His His His Asn Pro Ala Ser Ser Ile Asp Pro Asp
288      115      120      125
291 Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe
292      130      135      140

```

## RAW SEQUENCE LISTING

DATE: 07/19/2005

PATENT APPLICATION: US/10/541,513

TIME: 15:11:28

Input Set : A:\Final sequence list-12810-00106-US.txt

Output Set: N:\CRF4\07192005\J541513.raw

```

295 Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile
296 145          150          155          160
299 Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr
300          165          170          175
303 Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr
304          180          185          190
307 Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln
308          195          200          205
311 Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile
312          210          215          220
315 Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His
316 225          230          235          240
319 Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys
320          245          250
323 <210> SEQ ID NO: 5
324 <211> LENGTH: 1608
325 <212> TYPE: DNA
326 <213> ORGANISM: Haematococcus pluvialis
328 <220> FEATURE:
329 <221> NAME/KEY: CDS
330 <222> LOCATION: (3)..(971)
332 <400> SEQUENCE: 5
333 ct aca ttt cac aag ccc gtg agc ggt gca agc gct ctg ccc cac atc      47
334   Thr Phe His Lys Pro Val Ser Gly Ala Ser Ala Leu Pro His Ile
335   1          5          10          15
337 ggc cca cct cct cat ctc cat cgg tca ttt gct gct acc acg atg ctg      95
338 Gly Pro Pro Pro His Leu His Arg Ser Phe Ala Ala Thr Thr Met Leu
339          20          25          30
341 tcg aag ctg cag tca atc agc gtc aag gcc cgc cgc gtt gaa cta gcc      143
342 Ser Lys Leu Gln Ser Ile Ser Val Lys Ala Arg Arg Val Glu Leu Ala
343          35          40          45
345 cgc gac atc acg cgg ccc aaa gtc tgc ctg cat gct cag cgg tgc tcg      191
346 Arg Asp Ile Thr Arg Pro Lys Val Cys Leu His Ala Gln Arg Cys Ser
347          50          55          60
349 tta gtt cgg ctg cga gtg gca gca cca cag aca gag gag gcg ctg gga      239
350 Leu Val Arg Leu Arg Val Ala Ala Pro Gln Thr Glu Glu Ala Leu Gly
351          65          70          75
353 acc gtg cag gct gcc ggc gcg ggc gat gag cac agc gcc gat gta gca      287
354 Thr Val Gln Ala Ala Gly Ala Gly Asp Glu His Ser Ala Asp Val Ala
355 80          85          90          95
357 ctc cag cag ctt gac cgg gct atc gca gag cgt cgt gcc cgg cgc aaa      335
358 Leu Gln Gln Leu Asp Arg Ala Ile Ala Glu Arg Arg Ala Arg Arg Lys
359          100          105          110
361 cgg gag cag ctg tca tac cag gct gcc gcc att gca gca tca att ggc      383
362 Arg Glu Gln Leu Ser Tyr Gln Ala Ala Ala Ile Ala Ala Ser Ile Gly
363          115          120          125
365 gtg tca ggc att gcc atc ttc gcc acc tac ctg aga ttt gcc atg cac      431
366 Val Ser Gly Ile Ala Ile Phe Ala Thr Tyr Leu Arg Phe Ala Met His
367          130          135          140

```

**VERIFICATION SUMMARY**

DATE: 07/19/2005

PATENT APPLICATION: US/10/541,513

TIME: 15:11:29

Input Set : A:\Final sequence list-12810-00106-US.txt

Output Set: N:\CRF4\07192005\J541513.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application No

L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date